

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-124518

(43) 公開日 平成10年(1998) 5月15日

(51) Int.Cl.⁶

G 0 6 F 17/30

12/00

識別記号

5 4 5

5 4 7

F I

G 0 6 F 15/403

12/00

15/403

15/419

3 2 0 A

5 4 5 F

5 4 7 H

3 8 0 F

3 1 0

審査請求 有 請求項の数6 OL (全 11 頁)

(21) 出願番号

特願平8-274581

(22) 出願日

平成8年(1996)10月17日

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 山本 幸彦

東京都港区芝五丁目7番1号 日本電気株式会社内

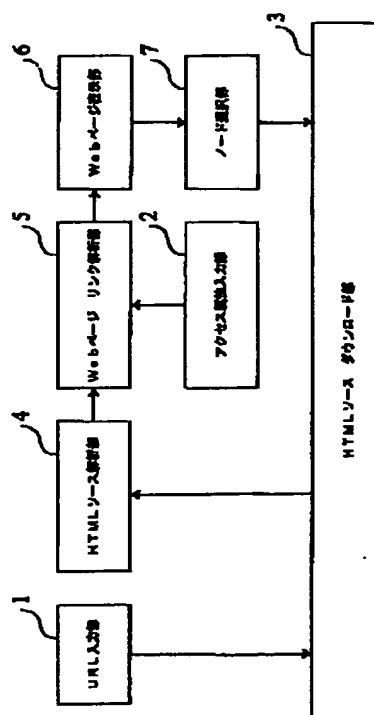
(74) 代理人 弁理士 後藤 洋介 (外2名)

(54) 【発明の名称】 優先アクセスウェブブラウジング処理装置

(57) 【要約】

【課題】 肥大化したWebページ及び階層の深いWebページでもユーザにとって分かりやすく、かつ簡単な方法で効率的にアクセスできるようにする。

【解決手段】 HTMLソースダウンロード部3では、URL(uniform resource locator)を指定して、WWW(World Wide Web)サーバからWebページ(HTML(Hyper Text Markup Language)ソース)を効率的にダウンロードを行う。Webページリンク解析部5では、リンクの張られたWebページのみを次々と自動的にアクセスする。そして、Webページ表示部6は、Webページの階層構造をツリー構造に変換して表示する。



【特許請求の範囲】

【請求項1】 ユニフォームリソースローケータ（URL）を入力するためのURL入力部と、アクセスしたいウェブ（Web）ページの属性を入力するアクセス属性入力部と、ワールドワイドウェブ（WWW）サーバからハイパーテキストマークアップ言語（HTML）ソースのダウンロードを行うHTMLソースダウンロード部と、該ダウンロードしたHTMLソースの解析を行うHTMLソース解析部と、前記Webページのリンク情報の解析を行うリンク解析部と、前記Webページの階層構造をツリー構造に変換して表示するWebページ表示部と、前記ツリー構造からノードを選択して該選択されたノードに対応するWebページをアクセスするノード選択部とを有することを特徴とする優先アクセスウェブブラウジング処理装置。

【請求項2】 請求項1に記載された優先アクセスウェブブラウジング処理装置において、前記リンク解析部は前記WWWサーバからダウンロードしたヘッダ情報を取り込む機能を備えるヘッダ情報取得部を有することを特徴とする優先アクセスウェブブラウジング処理装置。

【請求項3】 請求項1に記載された優先アクセスウェブブラウジング処理装置において、前記HTMLソース解析部は前記HTMLソースの中から代表とするテキスト情報を代表テキスト情報として決定する代表テキスト情報抽出部と、前記HTMLソースの中からハイパーリンク情報を抽出するハイパーリンク情報抽出部とを有することを特徴とする優先アクセスウェブブラウジング処理装置。

【請求項4】 請求項1に記載された優先アクセスウェブブラウジング処理装置において、前記Webページ表示部は前記Webページの階層構造をツリー構造に変換するWebページ階層構造／ツリー構造変換部を有することを特徴とする請求項1記載の優先アクセスウェブブラウジング処理装置。

【請求項5】 請求項1に記載された優先アクセスウェブブラウジング処理装置において、前記リンク解析部は前記Webページの属性のみを優先的にアクセスして前記Webページ相互のリンク情報を解析することを特徴とする優先アクセスウェブブラウジング処理装置。

【請求項6】 請求項1に記載された優先アクセスウェブブラウジング処理装置において、前記Webページ表示部は前記Webページの階層構造をツリー構造に変換して表示する際に、ノード内部に前記テキスト情報を表示するとともに前記Webページに含まれるタグ情報を用いてテキスト情報を抽出するようにしたこと特徴とする優先アクセスウェブブラウジング処理装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、相互にリンクの張られたウェブ（Web）ページ間の構成を高速に探索す

るアクセス方式に関し、特に、Webページの階層構造をツリー構造に変換する機能を有するWebブラウジング処理装置に関する。

【0002】

【従来の技術】 一般に、Webページ実現の手段として、ハイパーテキストが用いられており、ハイパーテキストを用いたブラウジング処理装置として、例えば、特開平4-321144号公報に記載された「ハイパーテキストのブラウジング処理装置」が知られている。

10 【0003】 このブラウジング処理装置では、必要な情報を入力する入力部と、入力された情報に所定の処理を施す入力処理部と、ハイパーテキスト作成に必要な編集操作を行うハイパーテキスト作成編集部と、グラフィカルに表示するために必要な処理を施すブラウジング処理部と、ブラウザ情報を表示するための表示処理部と、ブラウザの表示を行う表示部とを備えており、ハイパーテキスト相互間のリンク関係をネットワーク構造のものからツリー構造に変換して表示する機能を有している。

【0004】

20 【発明が解決しようとする課題】 ところで、上述の「ハイパーテキストのブラウジング処理装置」等の従来のWebページブラウジング処理装置では、ユーザはWebページ及びハイパーテキストを一つ一つマニュアル操作でアクセスしなければならず、このため、Webページ及びハイパーテキスト全体の構成を理解するためには、上位の階層から一つ一つ順に下の階層に向けてアクセスしなければならない。従って、Webページ及びハイパーテキスト全体の構成を理解する際には、極めて手間のかかる作業を行わなければならないという問題点がある。

30 【0005】 さらに、Webページ及びハイパーテキストには様々な種類のファイルがリンクされており、現状のWebブラウザ又はハイパーテキストブラウジング処理装置を用いた際、ユーザは必要な情報を含むWebページ又はハイパーテキストのブラウジングを行う場合に、ユーザが要求する情報以外のWebページ又はハイパーテキストであっても一通りアクセスしなければならない。このため、階層の深いWebページに対してアクセス時間がかかり、効率が悪いという問題点がある。

40 【0006】 特に、Webページがネットワークを介して存在する場合に、ネットワークの帯域幅が狭いと、サイズの大きなWebページ及び階層の深い位置にあるWebページをアクセスすると、ネットワーク構造をツリー構造に変換して表示する際に時間がかかってしまうという問題点がある。

【0007】 本発明の目的は、肥大化したWebページ及び階層の深い位置にあるWebページをアクセスする際にも効率的にアクセスすることのできる優先アクセスWebブラウジング処理装置を提供することにある。

50 【0008】 本発明の他の目的は、Webページ全体の

構成を容易に理解することのできるグラフィカルユーザインターフェースを実現可能な優先アクセスWebブラウジング処理装置を提供することにある。

【0009】

【課題を解決するための手段】本発明による優先アクセスWebブラウジング処理装置は、URL (uniform resource locator) を指定してWWW (WorldWide Web) サーバからWebページ (HTML (Hyper Text Markup Language) ソース) を効率的にダウンロードするHTMLソースダウンロード部と、リンクの張られたWebページのみを次々と自動的にアクセスするWebページリンク解析部と、Webページの階層構造をツリー構造に変換して表示するWebページ表示部とを備えることを特徴としている。そして、本発明では、ユーザは肥大化したWebページ及び階層の深いWebページでも効率的にアクセスすることができる。

【0010】さらに、Webページの階層構造をツリー構造に変換して表示しているから、Webページの全体構成の理解が容易となり、ユーザはツリー構造で表示されたノードを選択することによって、階層の深いWebページでも自在にアクセスすることができる。

【0011】また、ツリー構造のノード内部には、そのWebページを特徴づける情報を表示すれば、Webページを全てアクセスしなくても、その表示情報に基づいてWebページの内容を推測できる。

【0012】

【発明の実施の形態】以下本発明について図面を参照して説明する。

【0013】図1を参照して、図示の優先アクセスWebブラウジング処理装置は、URL {ユニフォームリソースローケータ (Uniform Resource Locator)} 入力部1、アクセス属性入力部2、HTMLソースダウンロード部3、HTMLソース {ハイパーテキストマークアップ言語 (HyperText Markup Language: HTML) ソース} 解析部4、Webページリンク解析部5、Webページ表示部6、及びノード選択部7を備えている。

【0014】URL入力部1では、アクセスしたいWebページ (HTMLソース) のURLを入力する。アクセス属性入力部2ではアクセスしたいWebページの属性を入力する。HTMLソースダウンロード部3は、ワールドワイドウェブ (World Wide Web: WWW) サーバからHTMLソースをダウンロードする。このダウンロードされたHTMLソースはHTMLソース解析部4によって解析される。Webページリンク解析部5では、Webページ相互のリンク情報を解析して、Webページ表示部6は、Webページの階層構造をツリー構造に変換する。そして、ノード選択部7はユーザのノードを選択に応じてノードに対応するWe

bページのアクセスを行う。

【0015】次に上述の各ブロックについて説明する。まず、図2を参照して、URL入力部について説明する。

【0016】URL入力部1はWebページURL入力部11、URL記録部12、及びURLデータベース13を備えており、WebページURL入力部11はユーザに対してURLの入力手段を提供する。つまり、ユーザはWebページURL入力部11によってURLを入力する。URL記録部12は、WebページURL入力部11から入力されたURLを受けこの入力URLをURLデータベース13に登録する。URLデータベース13には複数のURLを登録することができ、URLデータベース13を参照して、ユーザの入力したURLが取得される。

【0017】次に、図3を参照して、クセス属性入力部2について説明する。

【0018】アクセス属性入力部2はHTMLソースアクセス属性入力部21、アクセス属性データベース登録部22、及びアクセス属性データベース23を備えており、HTMLソースアクセス属性入力部21はユーザに対してアクセスしたいHTMLソースの属性を入力する手段を提供する。つまり、ユーザはHTMLソースアクセス属性入力部21を用いてアクセスしたいHTMLソースの属性を入力することになる。

【0019】アクセス属性データベース登録部22はHTMLソースアクセス属性入力部21から入力されたHTMLソースのアクセス属性をアクセス属性データベース23に登録する。このようにして、アクセス属性データベース23には、アクセス属性が登録され、アクセス属性データベース23を参照して、ユーザの入力したアクセス属性が取得される。

【0020】図4を参照して、HTMLソースダウンロード部3について説明する。

【0021】HTMLソースのダウンロード部3は、ファイル名/パス名参照部31及びGETメソッド実行部32を備えており、ファイル名/パス名参照部31にはURLデータベース13、ハイパーリンク情報データベース45、及びツリー表示データベース56が接続されている。

【0022】ファイル名/パス名参照部31はダウンロードするファイルの名前とパス名とを取得する。つまり、ファイル名/パス名参照部31はURLデータベース13を参照して、ユーザの入力したURLを取得する。ハイパーリンク情報データベース45及びツリー表示データベース56は、2回目以降のHTMLソースのダウンロードの際に参照されるが、ハイパーリンク情報データベース45及びツリー表示データベース56については後述する。

【0023】GETメソッド実行部32ではファイル名

／パス名参照部31で取得したファイルについてWWWサーバに対してGETメソッドを実行してダウンロードする。

【0024】図5を参照して、HTMLソース解析部4について説明する。

【0025】HTMLソース解析部はHTML解析部41、代表テキスト情報抽出部42、代表テキスト情報データベース43、ハイパーリンク情報抽出部44、及びハイパーリンク情報データベース45を備えており、HTML解析部41はHTMLソースの構文を解析して、HTMLソースに含まれるタグ情報及びテキスト情報を識別する。代表テキスト情報抽出部42はHTMLソースの中から代表とするテキスト情報を決定して、代表テキスト情報データベース43に登録する。

【0026】ハイパーリンク情報抽出部44はHTMLソースの中からハイパーリンク情報を抽出して、ハイパーリンク情報データベース45に登録する。

【0027】なお、代表テキスト情報データベース43とハイパーリンク情報データベース45を参照して、それぞれ代表テキスト情報、ハイパーリンク情報を取得することができる。

【0028】図6を参照して、Webページリンク解析部5について説明する。

【0029】Webページリンク解析部5は、ファイル名／パス名参照部51、HEADメソッド実行部52、ファイルヘッダ情報取得部53、ヘッダフィールド判定部54、ツリー表示データベース登録部55、及びツリー表示データベース56を備えており、ファイル名／パス名参照部51にはハイパーリンク情報データベース45が接続されている。また、ファイルヘッダ情報取得部53にはアクセス属性データベース23が接続されている。

【0030】ファイル名／パス名参照部51では、ハイパーリンク情報データベース45を参照して、リンクの張られたファイルの名前及びパス名を取得する。HEADメソッド実行部52ではファイル名／パス名参照部51で取得したファイルのヘッダ情報を取得するため、WWWサーバに対してHEADメソッドの実行を行う。

【0031】ファイルヘッダ情報取得部53はリンクの張られたファイルのヘッダ情報を取得する。つまり、ファイルヘッダ情報取得部53では、WWWサーバからダウンロードしたヘッダ情報を取り込む。

【0032】ヘッダフィールド判定部54では上述のようにして得られたヘッダ情報のContent-Type（コンテンツタイプ）の値が、アクセス属性データベース23に登録されているか否かを判定する。もし、ヘッダ情報のContent-Typeの値がアクセス属性データベース23に含まれているのであれば、ツリー表示データベース登録部55によって、ツリー表示データベース56にファイル名及びパス名が登録される。ヘ

ッダ情報のContent-Typeの値がアクセス属性データベース23に含まれていない場合には、ツリー表示データベース56には登録されない。

【0033】図7を参照して、Webページ表示部について説明する。

【0034】Webページ表示部6はツリー表示データベース参照部61、Webページ階層構造／ツリー構造変換部62、ツリー構造表示部63を備えており、ツリー表示データベース参照部61にはツリー表示データベース56が接続されている。

【0035】ツリー表示データベース参照部61ではツリー表示データベース56を参照して、ファイル名及びパス名を取得する。Webページ階層構造／ツリー構造変換部62は、ツリー表示データベース参照部61で取得したファイル名及びパス名の階層構造をツリー構造に変換する。ツリー構造表示部63は、Webページをツリー構造のノードとして描画し、階層構造をパスとして描画してツリー構造を表示する。

【0036】図8を参照して、ノード選択部について説明する。

【0037】ノード選択部7はノード指定部71及びファイル名／パス名決定部72を備えており、ファイル名／パス名決定部72にはツリー表示データベース56が接続されている。

【0038】ノード指定部71では、ユーザにツリー構造に含まれるノードの中から1つのノードを指定する手段を提供する。つまり、ユーザはノード指定部71を用いてツリー構造に含まれるNODEの中から一つのノードを指定する。ファイル名／パス名決定部72は、上述のようにして指定されたノードに対応するファイルの名前とパス名とを決定する。つまり、ファイル名／パス名決定部72では、ツリー表示データベース56を参照して、ノード指定部71で指定したノードに対応するファイル名とパス名の決定を行う。

【0039】次に、本発明による優先アクセスWebブラウジング処理装置の動作について説明する。

【0040】図9に示すように、クライアントはパソコン1aを用いてネットワークを介してWWWサーバ1bにアクセスする。クライアントがWWWサーバ1bに対して、HTMLソースの転送要求を行うと、WWWサーバ1bからHTMLソースがダウンロードされることになる。

【0041】図10を参照して、ユーザはURL入力部1を用いてアクセスしたいWebページを特定するURLを入力するとともにアクセス属性入力部2を用いてアクセスしたいWebページのアクセス属性を入力する（ステップS1）。具体的には、図11に示すように、まず、アクセスしたいWebページのURLを入力する（ステップS11）。そして、入力したURLをURL記録部12によって、URLデータベース13に登録す

る(ステップS12)。

【0042】次に、HTMLソースアクセス属性入力部21によって、ダウンロードするHTMLソースのアクセス属性を入力する(ステップS13)。そして、アクセス属性をアクセス属性データベース23に登録する(ステップS14)。

【0043】上述のようにして、ステップS1が終了すると、前述のように、HTMLソースダウンロード部3によって、WWWサーバからHTMLソースをダウンロードする(ステップS2)。具体的には、図12に示すように、URLデータベース13を参照して、ダウンロードするファイルの名前とパス名を取得する(ステップS21)。次に、GETメソッド実行部32によってGETメソッドを実行して、HTMLソースのダウンロードを行う(ステップS22)。

【0044】上述のようにして、ステップS2が終了すると、HTMLソース解析部4によって、HTMLソースの中から代表テキスト情報を抽出する(ステップS3)。具体的には、図13に示すように、まず、HTML解析部41によって、HTMLソースを解析して、HTMLソースに含まれるテキスト及びタグ情報を識別する(ステップS31)。

【0045】次に、WebページからTITLE(タイトル)タグを抽出する(ステップS32)。そしてTITLEタグが存在するか否かを判定する(ステップS33)。TITLEタグが存在する場合には、TITLEタグの情報をそのWebページの代表テキスト情報に決定して(ステップS34)。代表テキスト情報を代表テキスト情報データベース43に登録する(ステップS36)。

【0046】TITLEタグが存在しない場合には、そのWebページに含まれるテキスト情報の最初の一文を代表テキスト情報に決定して(ステップS35)、代表テキスト情報を代表テキスト情報データベース43に登録する(ステップS36)。

【0047】上述のようにして、ステップS3が終了すると、HTMLソース解析部4によって、さらにハイパーリンク情報が抽出される(ステップS4)。具体的には、図14に示すように、まず、HTML解析部41によって、HTMLソースを解析して(ステップS41)、Webページからタグ情報を抽出する(ステップS42)。次に、抽出したタグ情報がアンカータグ(ハイパーリンクを示すタグ)であるか否かを判定する(ステップS43)。

【0048】アンカータグであれば、タグ情報をハイパーリンク情報データベース45に登録する(ステップS44)。タグ情報がアンカータグでない場合は、ハイパーリンク情報データベース45には登録は行わず、ステップS42に戻る。

【0049】そして、ステップS42乃至S44を繰り返す

返して、そのHTMLソースに含まれる全てのアンカータグをハイパーリンク情報データベース45に登録する。

【0050】上述のようにして、ステップS4が終了すると、Webページリンク解析部5によって、HEADメソッドを実行してリンク先ファイルのヘッダフィールドを取得し、HTMLソースに含まれるファイルのリンク情報を解析する(ステップS5)。具体的には、図15に示すように、まず、ハイパーリンク情報データベース45を参照して、リンクの張られたファイルの名前とパス名を取得する(ステップS51)。

【0051】次に、HEADメソッド実行部52によって、HEADメソッドを実行して、リンクの張られたファイルのヘッダ情報を取得する(ステップS52)。ファイルヘッダ情報取得部53によって、取得したヘッダ情報の中からContent-Typeの値を抽出する(ステップS53)。

【0052】ヘッダフィールド判定部54によって、抽出したContent-Typeの値がtext/htmlであるか、もしくはアクセス属性データベース23に登録されているContent-Typeの値となっているかどうかを判定する(ステップS54)。

【0053】Content-Typeの値が上記の条件を満たせば、リンクの張られたファイルをツリー表示するものと決定して、ツリー表示データベース56に登録を行う(ステップS55)。Content-Typeの値が上記の条件を満たさなければ、ツリー表示データベース56への登録は行わない。

【0054】なお、ツリー表示データベース56に登録されたファイルは、前述のようにツリー構造のノードとして表示される。

【0055】その後、そのWebページにハイパーリンク情報が含まれているか否かを判定して(ステップS56)、含まれていれば上述のステップS52乃至S55を繰り返す。つまり、そのHTMLソースに含まれるリンク先ファイルの全てのヘッダ情報を調べて、ツリー表示のノードとして表示すべきファイルをツリー表示データベース56に登録する動作を繰り返す。

【0056】上述のようにして、ステップS5が終了すると、Webページの階層構造をツリー構造に変換してWebページの表示を行う(ステップS6)。具体的には、図16に示すように、ツリー表示データベース56を参照して、ツリー表示するファイル名、パス名、ファイルの個数、及びアクセス属性を取得する(ステップS61)。

【0057】ファイルの個数及びWebページの階層構造に応じて、ツリー構造のノードとして表示する際のノードの描画位置を決定する(ステップS62)。そして、ノードの描画を行い、代表テキスト情報データベース43を参照して得られた代表テキスト情報をノード内

部に表示する(ステップS63)。

【0058】次に、それぞれのノードをパスで接続してツリー構造を作成して表示する(ステップS64)。ユーザはツリー構造におけるノードを選択することによって、そのノードに対応するWebページをブラウジングすることができる。

【0059】さらにHTMLソースをダウンロードするか否かを判定して(ステップS7)、HTMLソース(Webページ)をダウンロードする場合には、ステップS2に戻る。そして、HTMLソース(Webページ)をダウンロードしない場合には、処理を終了する。

【0060】ここで、図17乃至図22を参照して、図17乃至図22にはWebページ(HTML)のソースの一例を示が示されており。ここでは、それぞれファイル名をa. html、b. html、c. html、d. html、e. html、f. htmlとして表す。

【0061】図23も参照して、図23には優先アクセスWebブラウジング処理装置で表示されたツリー構造の一例が示されている。なお、符号81乃至86で示す長方形のブロックはツリー構造のノードを表しており、各ノードはネットワーク上で相互にリンクされているWebページに対応している。前述のように、ユーザはノードを選択することによって、そのノードに対応したWebページをブラウジングすることができる。

【0062】図示のツリー構造では、ノード81の下位にノード82とノード86とが直接リンクで張られており、ノード82の下位にはノード83とノード84とがリンクされている。そして、ノード84の下位にはノード85がリンクされている。なお、ノード83、ノード85、及びノード86には、下位のリンクは存在しない。

【0063】各ノードの内部には代表テキスト情報が表示される。代表テキスト情報はそのWebページの内容を示すキーワードであり、ユーザはこのテキスト情報から、そのWebページの内容をある程度推測することができる。従って、ユーザは効率的なWebページのアクセスを行うことができる。

【0064】図17は、ノード81のWebページ(a. html)の内容を示しており、このWebページにはTITLEタグが含まれている。従って、代表テキスト情報抽出部42によって、このWebページの代表テキスト情報は「インデックス」となる。よって、ノード81の内部には「インデックス」が表示される。

【0065】また、このWebページにはアンカタグが存在し、アンカタグが表すファイル名はそれぞれb. html、f. htmlとなる。よって、ノード81の下位にはb. htmlを表すノード82とf. htmlを表すノード86がリンクされている。

【0066】図18は、ノード82のWebページ

(b. html)の内容を示しており、このWebページにはTITLEタグが含まれている。従って、このWebページの代表テキスト情報は「第2階層」となる。よって、ノード82の内部には「第2階層」と表示される。

【0067】また、このWebページにはアンカタグが存在し、アンカタグが表すファイル名はそれぞれc. htmlとd. htmlになる。よって、ノード82の下位にはc. htmlを表すノード83とd. htmlを表すノード84がリンクされている。

【0068】図19はノード83のWebページ(c. html)の内容を示しており、このWebページにはTITLEタグが含まれていない。そこで、代表テキスト情報抽出部42によって、テキスト情報の最初の一文を抽出する。これによって、このWebページの代表テキスト情報は「第3階層です。」となる。よって、ノード83の内部には「第3階層です。」と表示される。なお、このWebページにはハイパーリンク情報が存在しないので、ノード83から下位へ向かうパスは存在しない。

【0069】図20はノード84のWebページ(d. html)の内容を示しており、このWebページにはTITLEタグが含まれている。代表テキスト情報抽出部42によって、このWebページの代表テキスト情報は「第3階層」となる。よって、ノード84の内部には「第3階層」と表示される。

【0070】なお、このWebページにはアンカタグが存在し、アンカタグが表すファイル名はe. htmlとなる。よって、ノード84の下位にはe. htmlを表すノード85がリンクされる。

【0071】図21は、ノード85のWebページ(e. html)の内容を示しており、このWebページにはTITLEタグが含まれている。従って、代表テキスト情報抽出部42によって、このWebページの代表テキスト情報は「第4階層」となる。よって、ノード85の内部には「第4階層」と表示される。なお、このWebページにはハイパーリンク情報が存在しないので、ノード83と同様に、下位へ向かうパスは存在しない。

【0072】図22はノード86のWebページ(f. html)の内容を示しており、WebページにはTITLEタグが含まれていない。そこで、代表テキスト情報抽出部42によって、このWebページの代表テキスト情報は「第2階層です。」となる。よって、ノード86の内部には「第2階層です。」と表示される。なお、このWebページにはハイパーリンク情報が存在しないので、ノード83及びノード85と同様に、下位へ向かうパスは存在しない。

【0073】このように、本発明の優先アクセスWebブラウジング処理装置によって、ユーザは肥大化したW

10

20

30

40

50

Webページ及び階層の深いWebページでも、ツリー構造を用いたユーザインターフェースとヘッダフィールドを用いたリンク解析アルゴリズムによって、効率的なWebページのブラウジングを行うことができる。

【0074】さらに、ツリー構造を用いたユーザインターフェースによって、Webページ全体の構成の理解も容易になる。

【0075】

【発明の効果】以上説明したように、本発明では、肥大化したWebページ及び階層の深いWebページをブラウジングする際にも、Webページ及び階層の深いWebページを効率的にアクセスすることができるという効果がある。

【0076】さらに、本発明では、Webページの階層構造をツリー構造に変換して表示するグラフィカルユーザインターフェースを用いているから、Webページの全体構成の理解が容易となる。その結果、ユーザはツリー構造で表示されたノードを選択するだけで、階層の深いWebページでも自在にアクセスすることができるという効果がある。

【図面の簡単な説明】

【図1】本発明による優先アクセスWebブラウジング処理装置の一例を示す図である。

【図2】図1に示すURL入力部を詳細に示すブロック図である。

【図3】図1に示すアクセス属性入力部を詳細に示すブロック図である。

【図4】図1に示すHTMLソースダウンロード部を詳細に示すブロック図である。

【図5】図1に示すHTMLソース解析部を詳細に示すブロック図である。

【図6】図1に示すWebページリンク解析部を詳細に示すブロック図である。

【図7】図1に示すWebページ表示部を詳細に示すブロック図である。

【図8】図1に示すノード選択部を詳細に示すブロック

図である。

【図9】WWWサーバからHTMLソースをダウンロードする際の動作を示す図である。

【図10】本発明による優先アクセスWebブラウジング処理装置の動作を説明するための流れ図である。

【図11】URL及びアクセス属性を入力する際の動作を説明するための流れ図である。

【図12】WWWサーバからHTMLソースをダウンロードする際の動作を説明するための流れ図である。

【図13】代表テキスト情報を決定する際の動作を説明するための流れ図である。

【図14】ハイパーリンク情報を抽出する際の動作を説明するための流れ図である。

【図15】リンク先ファイルのヘッダフィールドを取得してリンク情報を解析する際の動作を説明するための流れ図である。

【図16】Webページの階層構造をツリー構造に変換して表示する際の動作を説明するための流れ図である。

【図17】Webページソースの一例を表す図である。

【図18】Webページソースの一例を表す図である。

【図19】Webページソースの一例を表す図である。

【図20】Webページソースの一例を表す図である。

【図21】Webページソースの一例を表す図である。

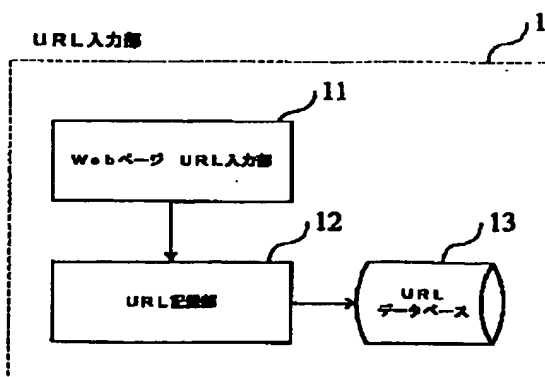
【図22】Webページソースの一例を表す図である。

【図23】本発明による優先アクセスWebブラウジング処理装置を用いて表示したツリー構造の一例を示す図である。

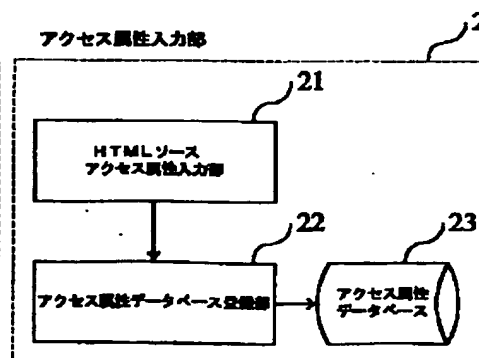
【符号の説明】

- 1 URL入力部
- 2 アクセス属性入力部
- 3 HTMLソースダウンロード部
- 4 HTMLソース解析部
- 5 Webページリンク解析部
- 6 Webページ表示部
- 7 ノード選択部

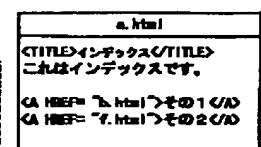
【図2】



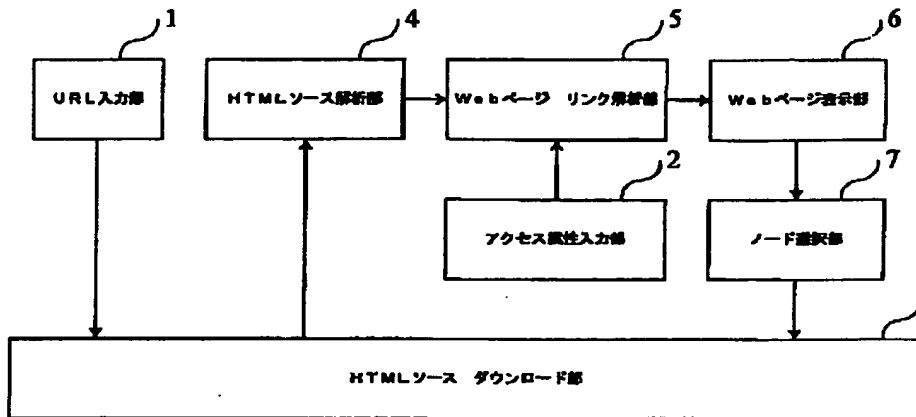
【図3】



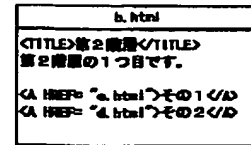
【図17】



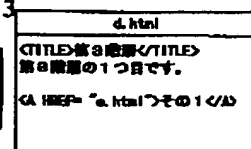
【図1】



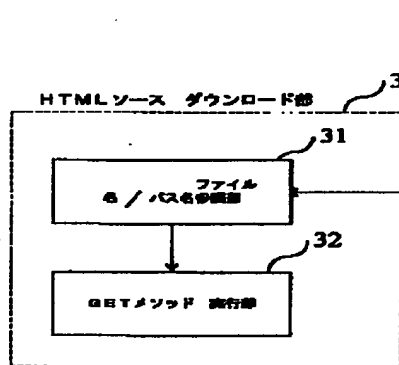
【図18】



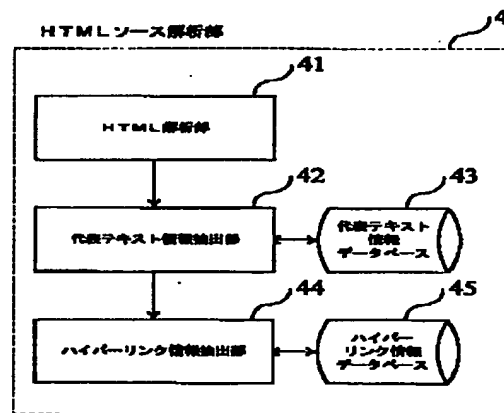
【図20】



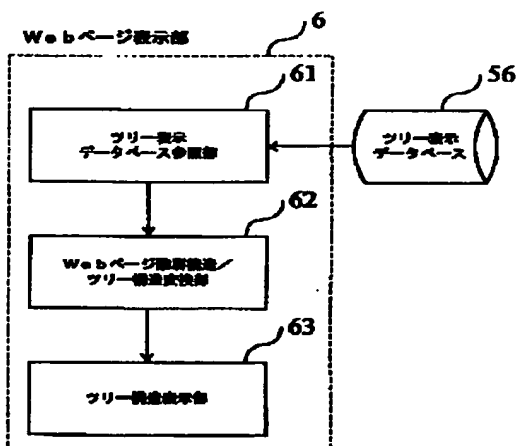
【図4】



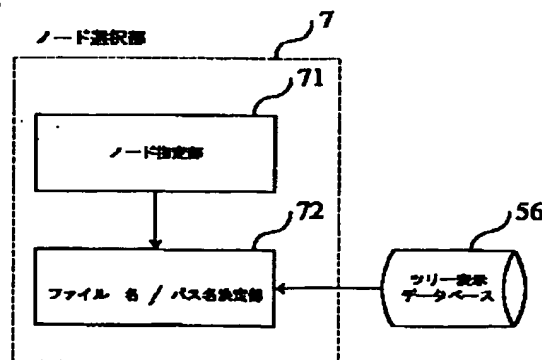
【図5】



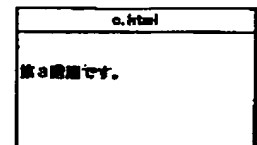
【図7】



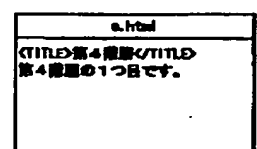
【図8】



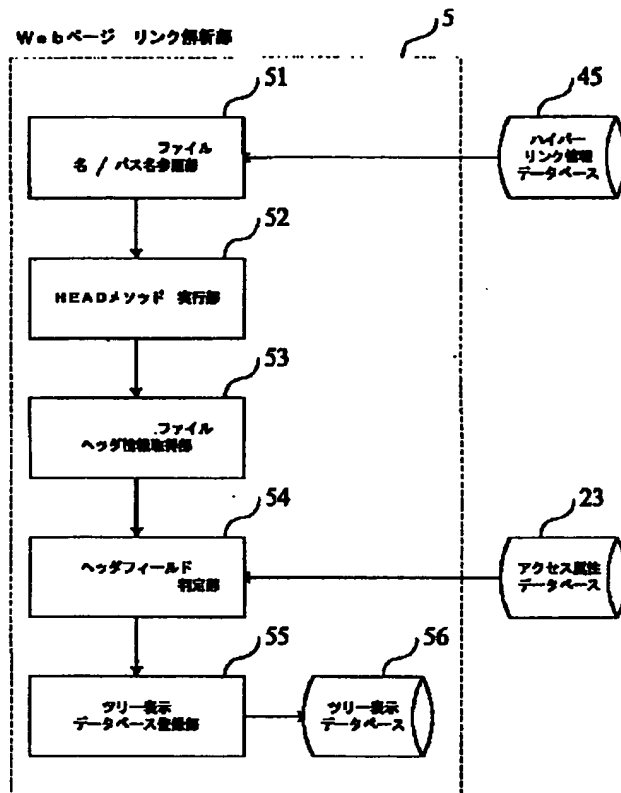
【図19】



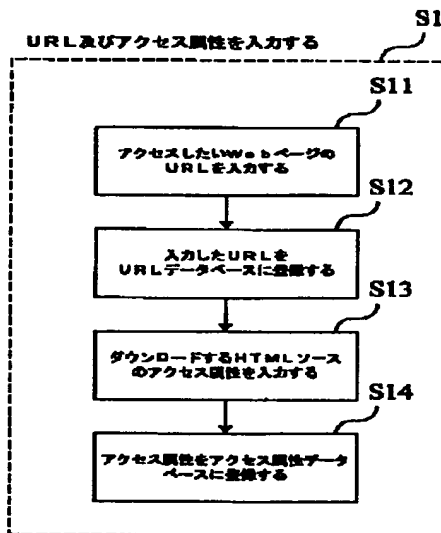
【図21】



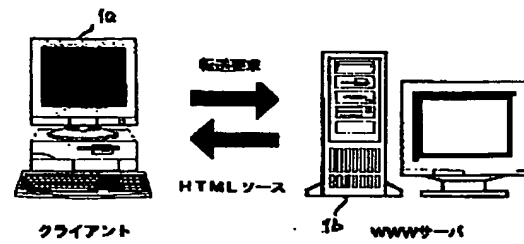
【図6】



【図11】

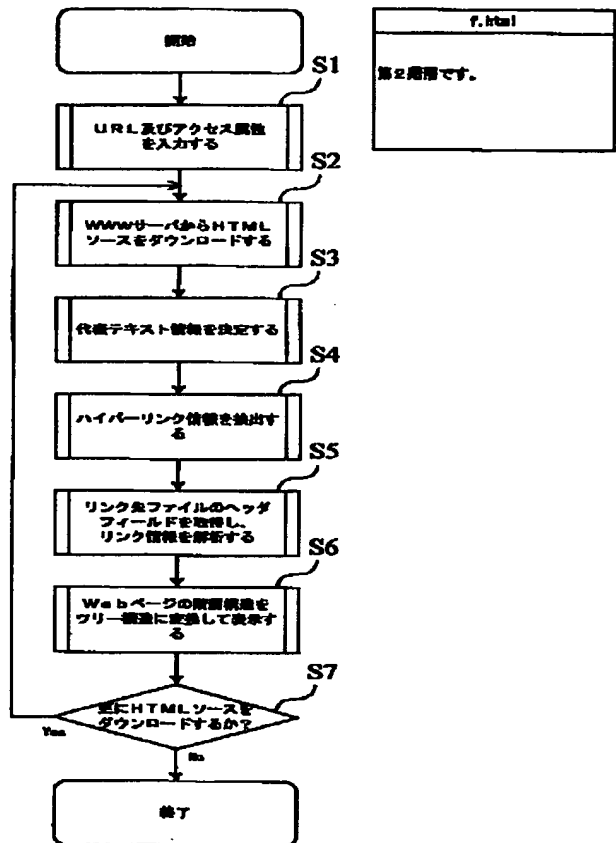


【図9】

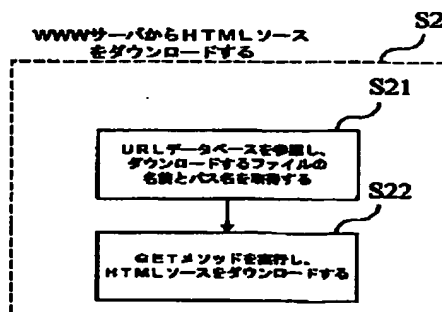


【図10】

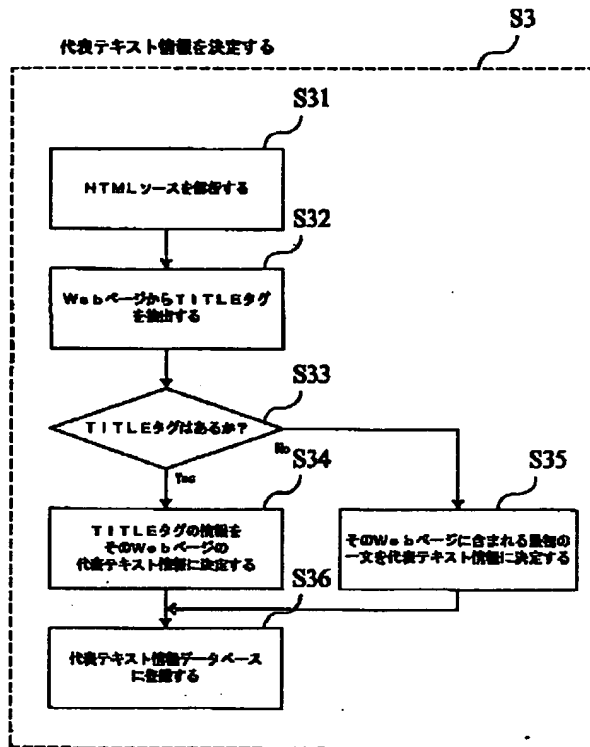
【図22】



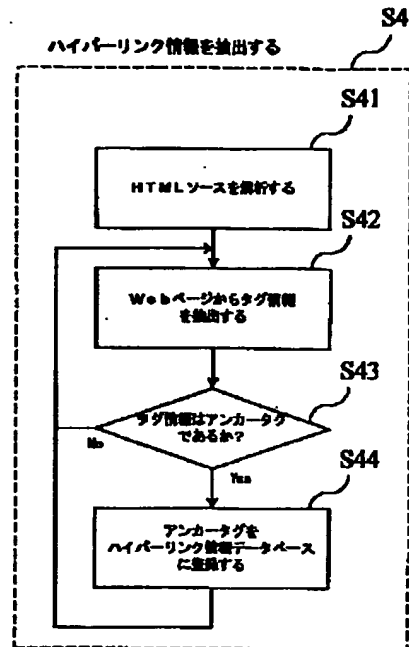
【図12】



【図13】

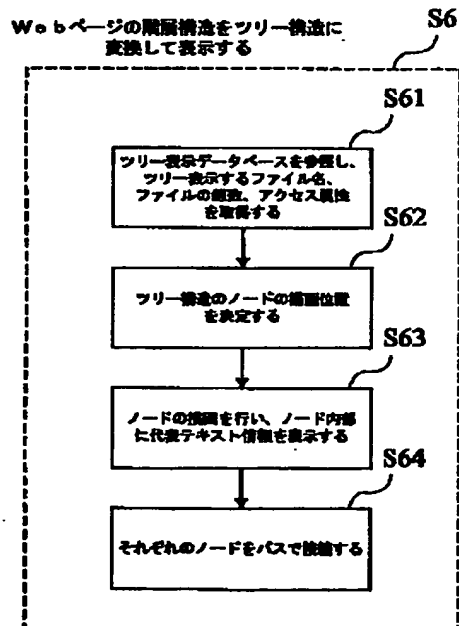


【図14】

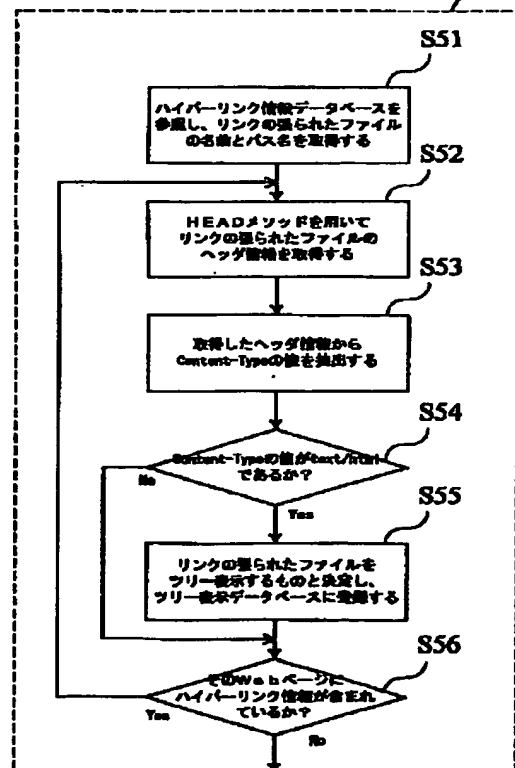


【図15】

【図16】



リンク先ファイルのヘッダフィールドを取得し、リンク情報を解析する S5



【図23】

